Remarks

The above-referenced application has been reviewed in light of the Examiner's Office Action dated December 12, 2003. Replacement Figures 1, 2a, and 2b have been included, in response to objections by the Draftspersons. The Abstract of the Disclosure has been amended, per the Examiner's recommendations. Claim 13 has been amended, per Examiner's recommendations to make the claim allowable. Accordingly, Claims 1-22 are currently pending in this application. No New Matter has been added. The Examiner's reconsideration of the rejections in view of the following remarks is respectfully requested.

In accordance with the Office Action, Claims 1-12 and 14-22 stands rejected under 35 U.S.C. §102(b), as being anticipated by Rosenbaum et al ('Rosenbaum hereinafter), US Patent 4,384,329.

With respect to independent Claim 1, applicants respectfully submit with traverse that Rosenbaum fails to disclose each and every element, as set forth in independent Claim 1 of the current application, either expressly, or inherently, for at least the reasons set forth herein.

The Office Action cites Rosenbaum to teach, (in col. 1, line 62 to col. 3, line 2 and Fig. 2),"'the value of each element V_{ij} of index matrix V is a function of the number of occurrences of the i^{th} vocabulary term in the j^{th} document' as synonym and antonym data base structures and text processing systems control for interactively accessing these data base structures is implemented by devising a symmetric binary matrix storage

organization which creates a word-wise relational data base linking the respective entries in a word list for retrieval while using minimum storage and without incurring entry redundancy…." Applicants respectfully disagree. The Rosenbaum matrix is a square binary matrix wherein, "Each row has a binary "1" bit set in the column position corresponding to each other word that is a synonym or antonym for the word defined by the row. All other positions in the row are set to binary '0'" (Rosenbaum, column 2, lines 10-14). In contrast, each element v_{ij} of rectangle index matrix V, according to current application, is a natural number representing the number of occurrences a particular word appears in a document. Each element v_{ij}, cannot be represented by a single bit binary number, as shown in Rosenbaum's matrix. Thus, Rosenbaum does not disclose, "the value of each element v_{ij} of index matrix V is a function of the number of occurrences of the i^{th} vocabulary term in the j^{th} document", as recited in Claim 1 of the present application.

The Office Action cites Rosenbaum to teach, (in col. 2, lines 15-17 et. seq.), "'factoring out non-negative matrix factors T and D such that $V \approx TD$' as overall size of the matrix is reduced by run-length encoding the number of column position 1 bits in each row…." and (in col. 2, lines 2-11 and col. 4, lines 55-64, Fig. 2), "wherein T is an n x r term matrix, D is an r x m document matrix, and r<nm/(n+m)"

Applicants respectfully disagree. Rosenbaum does not disclose "factoring out non-negative matrix factors T and D such that $V \approx TD$". The run-length encoding used to reduce the size of the Rosenbaum matrix does so without loosing any data. The non-negative matrix factorization, described in the current application, is fundamentally different from run-length encoding, as described in Rosenbaum. Non-negative matrix

factorization is a lossy process, one where, a tradeoff must be made between the compressing the matrix and avoiding data loss. Thus, Rosenbaum does not disclose a matrix size reduction, as to result in data loss, as such, Rosenbaum does not disclose, "factoring out non-negative matrix factors T and D such that $V \approx TD$", as claimed in Claim 1 of the present application. Further, since the Rosenbaum matrix is loss less, the matrix may have a rank factorization r where 'r=nm/(n+m)', but not a rank factorization where "r<nm/<n+m)", as claimed in Claim 1 of the present application.

Therefore, applicants respectfully submit that Claim 1 is not anticipated nor rendered obvious by Rosenbaum.

With respect to independent Claim 14, applicants respectfully submit with traverse that Rosenbaum fails to disclose each and every element, as set forth in independent Claim 14 of the current application, either expressly, or inherently, for at least the reasons given herein.

The Office Action cites Rosenbaum to teach, (in col. 5, lines 1-15), "the value of each element v_{ij} of index matrix V is a function of the number of occurrences of the i^{th} vocabulary term in the j^{th} document". Applicants respectfully disagree. In column 5, lines 1-15, Rosenbaum discloses finding the synonym of a vocabulary term, but does not disclose, "the value of each element v_{ij} of index matrix V is a function of the number of occurrences of the i^{th} vocabulary term in the j^{th} document". Also, as discussed earlier, each element v_{ij}, cannot be represented by a single bit binary number, as shown in Rosenbaum's matrix. Thus, Rosenbaum does not disclose, "the value of each element v_{ij}

of index matrix V is a function of the number of occurrences of the i^{th} vocabulary term in the j^{th} document", as recited in Claim 14 of the present application.

The Office Action cites Rosenbaum to teach, (in the Abstract et seq), "factoring out non-negative matrix factors T and D such that $V \approx TD$" and (in the col. 5, lines 1-15 et seq), "wherein T is an $n \times r$ term matrix, D is an $r \times n$ document matrix, and $r<nm/(n+m)$"

Applicants respectfully disagree. As discussed above, Rosenbaum not disclose "factoring out non-negative matrix factors T and D such that $V \approx TD$". The run-length encoding used to reduce the size of the Rosenbaum matrix does so without loosing any data. The non-negative matrix factorization, described in the current application, is fundamentally different from run-length encoding, as described in Rosenbaum. Non-negative matrix factorization is a lossy process, one where, a tradeoff must be made between the compressing the matrix and avoiding data loss. Thus, Rosenbaum does not disclose a matrix size reduction, as to result in data loss, as such, Rosenbaum does not disclose, "factoring out non-negative matrix factors T and D such that V≈TD", as claimed in Claim 14 of the present application. Further, since the Rosenbaum matrix is loss less, the matrix may have a rank factorization r where 'r=nm/(n+m)', but not a rank factorization where "r<nm/<n+m)", as claimed in Claim 14 of the present application.

Therefore, applicants respectfully submit that Claim 14 is not anticipated nor rendered obvious by Rosenbaum.

With respect to independent Claim 15, applicants respectfully submit with traverse that Rosenbaum fails to disclose each and every element, as set forth in

independent Claim 15 of the current application, either expressly, or inherently, for at least the reasons given herein.

The Office Action cites Rosenbaum to teach, (in col. 2, lines 2-4), "'an $r \times m$ document matrix D, such that $V \approx TD$ wherein T is an $n \times r$ term matrix' as data base structure devising a NXN binary matrix where N is equal to the number of words in the vocabulary." and (in col. 5, lines 55-67 et seq), "V is a non-negative $n \times m$ index matrix, wherein each of its m columns represents an j^{th} document having n entries containing the value of a function of the number of occurrences of a i^{th} term appearing in said j^{th} document; and wherein T and D are non-negative matrix factors of V and $r<nm/(n+m)$"

Applicants respectfully disagree. Rosenbaum does not disclose, "an $r \times m$ document matrix D, such that $V \approx TD$ wherein T is an $n \times r$ term matrix", nor does it disclose, "wherein T and D are non-negative matrix factors of V and $r<nm/(n+m)$", both of which are recited in Claim 15 of the present application. Thus, since both elements are recited in Claim 15 of the present application, they can be said to describe a document matrix D and the term matrix T, which makes up "$V \approx TD$", where T and D are "non-negative matrix factors of V". As discussed above, the non-negative matrix factorization, described in the current application, is fundamentally different from run-length encoding, as described in Rosenbaum. Non-negative matrix factorization is a lossy process, one where, a tradeoff must be made between the compressing the matrix and avoiding data loss. Thus, Rosenbaum does not disclose a matrix size reduction, as to result in data loss, as such, Rosenbaum does not disclose, "an $r \times m$ document matrix D, such that $V \approx TD$ wherein T is an $n \times r$ term matrix", as claimed in Claim 15 of the present application. Further, since the Rosenbaum matrix is loss less, the matrix may have a rank factorization

r where 'r=nm/(n+m)', but not a rank factorization where "r<nm/<n+m)", as claimed in Claim 15 of the present application.

Further, still, Claim 15 of the present application recites, "*V* is a non-negative *n x m* index matrix, wherein each of its *m* columns represents an j^{th} document having *n* entries containing the value of a function of the number of occurrences of a i^{th} term appearing in said j^{th} document". Each element in matrix index V can be described by v_{ij}, where *i* represents the "i^{th} term" and j represents the "j^{th} document". As discussed above, each element v_{ij} of index matrix *V*, according to the current application, is a natural number representing the number of occurrences a particular word appears in a document. Each element v_{ij}, cannot be represented by a single bit binary number, as shown in Rosenbaum's NxN vocabulary matrix. Thus, Rosenbaum does not disclose, "*V* is a non-negative *n x m* index matrix, wherein each of its *m* columns represents an j^{th} document having *n* entries containing the value of a function of the number of occurrences of a i^{th} term appearing in said j^{th} document", as recited in Claim 15 of the present application.

Therefore, applicants respectfully submit that Claim 15 is not anticipated nor rendered obvious by Rosenbaum.

With respect to independent Claims 16 and 22, applicants respectfully submit, with traverse, that Rosenbaum fails to teach each and every element, as set forth in the claims, either expressly, or inherently, for at least the reasons given herein.

The Office Action cites Rosenbaum to teach, (e.g., col. 1, line 62 to col. 3, line 2 and Fig. 2), "'performing a first pass retrieval through a first database representation and scoring m retrieved documents according to relevance to said query' as synonym and

antonym...redundancy." Applicants respectfully disagree. In, column 2, lines 23-24, Rosenbaum states, "The accessed words are output by the control system and displayed to an operator." By stating that the results are displayed to an operator, without mentioning any form of ranking of the results, Rosenbaum teaches away from what is claimed in Claims 16 and 22 of the current application. Thus, Rosenbaum does not disclose the element, "performing a first pass retrieval through a first database representation and scoring m retrieved documents according to relevance to said query", as recited in Claim 16 and 22 of the current application.

The Office Action cites Rosenbaum to teach, (e.g., col. 5, lines 1-50 et seq), "executing a second pass retrieval through a second database representation and scoring documents retrieved from said first pass retrieval so as to generate a final relevancy score for each document". Applicants respectfully disagree. Rosenbaum does not disclose, nor imply, " generate a final relevancy score for each document" as recited in Claims 16 and 22 of the current application. As discussed earlier, Rosenbaum teaches away from the generation of such a score. Thus, Rosenbaum does not disclose the element, "executing a second pass retrieval through a second database representation and scoring documents retrieved from said first pass retrieval so as to generate a final relevancy score for each document", as recited in Claim 16 and 22 of the current application.

The Office Action cites Rosenbaum to teach, (e.g., col. 5, lines 55-67 et seq), "wherein said second database representation comprises an $r \times m$ document matrix D, such that $V \approx TD$ wherein T is an $n \times r$ term matrix", and (e.g., col. 2, lines 2-4 and col. 5 lines 1-15), "wherein T and D are non-negative matrix factors of V and $r<nm/(n+m)$".

Applicants respectfully disagree. Rosenbaum does not disclose, "wherein said second database representation comprises an r x m document matrix D, such that V≈TD wherein T is an n x r term matrix", nor does it disclose, "wherein T and D are non-negative matrix factors of V and r<nm/(n+m)", as recited in both Claims 16 and 22. Since both elements are recited in Claims 16 and 22, they can be said to describe a document matrix D and the term matrix T, which makes up "V≈TD", where T and D are "non-negative matrix factors of V". As discussed above, the non-negative matrix factorization, described in the current application, is fundamentally different from run-length encoding, as described in Rosenbaum. Non-negative matrix factorization is a lossy process, one where, a tradeoff must be made between the compressing the matrix and avoiding data loss. Thus, Rosenbaum does not disclose a matrix size reduction, as to result in data loss, as such, Rosenbaum does not disclose, "wherein said second database representation comprises an r x m document matrix D, such that V≈TD wherein T is an n x r term matrix", as claimed in both Claims 16 and 22 of the present application. Further, since, as discussed above, the Rosenbaum matrix is loss less, the matrix may have a rank factorization r where 'r=nm/(n+m)', but not a rank factorization where "r<nm/<n+m)", as claimed in both Claims 16 and 22 of the present application.

The Office Action cites Rosenbaum to teach, (e.g., col. 5, lines 1-15, and Figs. 2-3), "V is a non-negative *n x m* index matrix, wherein each of its m columns represents an j^{th} document having *n* entries containing the value of a function of the number of occurrences of a i^{th} term of said vocabulary appearing in said j^{th} document". Applicants respectfully disagree. Each element in matrix index V can be described by v_{ij}, where *i* represents the "i^{th} term" and j represents the "j^{th} document". As discussed above, each

element v_{ij} of index matrix V, according to the current application, is a natural number representing the number of occurrences a particular word appears in a document. Each element v_{ij}, cannot be represented by a single bit binary number, as shown in Rosenbaum's NxN vocabulary matrix. Thus, Rosenbaum does not disclose, "V is a non-negative n x m index matrix, wherein each of its m columns represents an j^{th} document having n entries containing the value of a function of the number of occurrences of a i^{th} term appearing in said j^{th} document", as recited in both Claims 16 and 22 of the present application.

Therefore, applicants respectfully submit that both Claims 16 and 22 are not anticipated nor rendered obvious by Rosenbaum.

With respect to dependent Claims 2-12 and 17-21, applicants' respectfully submit with traverse that Rosenbaum, as cited in the current Office Action, fails to disclose each and every element, as set forth in the dependent claims of the current application, either expressly, or inherently, for at least the reasons set forth herein. Claims 1 and 16 are not disclosed by Rosenbaum. Claims 2-12 depend from Claim 1 and Claims 17-21 depend from claim 16, necessarily including each of the elements and limitations thereof. Thus, each and every element of Claims 2-12 and 17-21 are not taught by Rosenbaum.

The cited references do not teach, either alone or combined, what is claimed by the current Claims of the present application. Therefore, it is respectfully submitted that Claims 1, amended 13, 14, 15, 16, and 22, and those that depend therefrom, are patentable over the cited Rosenbaum reference, as well as all other references of record in this case.

Conclusion

Accordingly, it is respectfully submitted that independent Claims 1, amended 13,14, 15, 16, and 22 are in condition of allowance for at least the reasons stated above. Since Claims 2-13 and 17-21 each depend from one of the above claims and necessarily include each of the elements and limitations thereof, it is respectfully submitted that these claims are also in condition for allowance for at least the reasons stated. Thus, each of Claims 1-22 is in condition for allowance. All issues raised by the Examiner having been addressed, reconsideration of the rejections and an early and favorable allowance of this case is earnestly solicited.

For the foregoing reasons, the present application including claims 1-22 is believed to be in condition for allowance. The Examiner's early and favorable action is respectfully urged.

Respectfully submitted,

Frank Chau
Reg. No. 34,316
Aniket A. Patel
Reg. No. 55,525
Attorney for Applicant(s)

F. Chau & Associates, LLC
1900 Hempstead Tnpk., Suite 501
East Meadow, NY 11553
TEL.: (516) 357-0091
FAX: (516) 357-0092